ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

Semantic Image Segmentation with Contextual Hierarchical Models

K.Saraswathi[1], C.Tamilselvi[2]

M. Phil Scholar, Department of Computer Science and Research, Bharathiyar Arts and Science College (w), Deviyakurichi , Attur (tk),Salem(dt) , Tamil Nadu, India[1]

Assistant Professor, Department of Computer Science and Research, Bharathiyar Arts and Science College (w), Deviyakurichi , Attur (tk),Salem(dt), Tamil Nadu, India[2]

ABSTRACT: Semantic segmentation is the problem of assigning an object label to each pixel. It unifies the image segmentation and object recognition problems. The importance of using contextual information in semantic segmentation frameworks has been widely realized in the field. We propose a contextual framework, called contextual hierarchical model (CHM), which learns contextual information in a hierarchical framework for semantic segmentation. At each level of the hierarchy, a classifier is trained based on down sampled input images and outputs of previous levels. Our model then incorporates the resulting mustier solution contextual information into a classifier to segment the input image at original resolution. This training strategy allows for optimization of a joint posterior probability at multiple resolutions through the hierarchy. Contextual hierarchical model is purely based on the input image patches and does not make use of any fragments or shape examples. Hence, it is applicable to a variety of problems such as object segmentation and edge detection. We demonstrate that CHM outperforms state-of-the-art on Stanford background and Weizmann horse datasets. It also outperforms state-of-the-art edge detection methods on NYU depth dataset and achieves state-of-the-art on Berkeley segmentation dataset.

I.INTRODUCTION

The human (primate) visual system observes and makes sense of a dynamic scene (video) or a static scene (image) by making a series of fixations at various salient locations in the scene. The eye movement between consecutive fixation called a saccade. Even during a fixation, the human eye is continuously moving. Such movement is called fixation movement.

The main distinction between the fixation eye movements during a fixation and saccades between fixations is that the former is an involuntary movement whereas the latter is a voluntary Movement.

But the important question is: Why does the human visual system make these eye movements? One obvious role of fixations—the voluntary eye movements—is capturing high resolution visual information from the salient locations in the scene as the structure of the human retina has a high concentration of cones (with fine resolution) in the central fovea.

However, psychophysics suggests a more critical role of fixations in visual perception. For instance, during a change blindness experiment, the subjects were found to be unable to notice a change when their eyes were fixated at a location away from where the change had occurred in the scene unless the change altered the gist or the meaning of the scene.

In contrast, the change is detected quickly when the subjects fixated on the changing stimulus or close to it. This clearly suggests a more fundamental role of fixation in how we perceive a scene (or image).

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

II.PROJECT DESCRIPTION

A. EXISTING SYSTEM

The importance of using contextual information in semantic segmentation frameworks has been widely realized in the field. We existing a contextual framework, called contextual hierarchical model (CHM), which learns contextual information in a hierarchical framework for semantic segmentation. At each level of the hierarchy, a classifier is trained based on down sampled input images and outputs of previous levels. Our model then incorporates the resulting multiresolution contextual information into a classifier to segment the input image at original resolution.

This training strategy allows for optimization of a joint posterior probability at multiple resolutions through the hierarchy. Contextual hierarchical model is purely based on the input image patches and does not make use of any fragments or shape . Hence, it is applicable to a variety of problems such as object segmentation and edge detection. We demonstrate that CHM outperforms state-of-the-art on Stanford background and Weizmann horse datasets.

2.1.1 DISADVANTAGES

- Segmentation frameworks have been widely realized in the field.
- We existing a contextual framework, called contextual hierarchical model
- At each level of the hierarchy, a classifier is trained based on down sampled input images and outputs of previous levels.

B. PROPOSED SYSTEM

The proposed segmentation process is carried out in two separate steps: First, all visual cues are combined to generate the probabilistic boundary edge map of the scene; second, in this edge map, the “optimal” closed contour around a given fixation point is found. Having two separate steps also makes it possible to establish a simple feedback between the mid-level cue (regions) and the low-level visual cues (edges). In fact, we propose a segmentation refinement process based on such a feedback process.

Finally, our experiments show the promise of the proposed method as an automatic segmentation framework for a general purpose visual system.

2.2.1 ADVANTAGES

- Visual cues are combined to generate the probabilistic boundary edge map of the scene.
- Having two separate steps also makes it possible to establish a simple feedback.
- Optimal” closed contour around a given fixation point is found.
- We propose a segmentation refinement process based on such a feedback process.
- Proposed method as an automatic segmentation framework for a general purpose visual system.

III. LITERATURE SURVEY

A. OVERVIEW

Besides enlarging your knowledge about the topic, writing a literature review lets you gain and demonstrate skills in two areas.

1. INFORMATION SEEKING: The ability to scan the literature efficiently, using manual or computerized methods, to identify a set of useful articles and books.

2. CRITICAL APPRAISAL: The ability to apply principles of analysis to identify unbiased and valid A literature review is an account of what has been published on a topic by accredited scholars and researchers. Occasionally you will be asked to write one as a separate assignment, but more often it is part of the introduction to an essay, research report, or thesis. In writing the literature review, your purpose is to convey to your reader what knowledge and ideas have been established on a topic, and what their strengths and weaknesses are.

As a piece of writing, the literature review must be defined by a guiding concept (e.g., your research objective, the problem or issue you are discussing or your argumentative thesis). It is not just a descriptive list of the studies.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

3.1 M. SEYEDHOSSEINI AND T. TASDIZEN, "MULTI-CLASS MULTI-SCALE SERIES CONTEXTUAL MODEL FOR IMAGE SEGMENTATION,"

Contextual information has been widely used as a rich source of information to segment multiple objects in an image. A contextual model uses the relationships between the objects in a scene to facilitate object detection and segmentation. Using contextual information from different objects in an effective way for object segmentation, however, remains a difficult problem. In this paper, we introduce a novel framework, called multiclass multiscale (MCMS) series contextual model, which uses contextual information from multiple objects and at different scales for learning discriminative models in a supervised setting. The MCMS model incorporates cross-object and inter-object information into one probabilistic framework and thus is able to capture geometrical relationships and dependencies among multiple objects in addition to local information from each single object present in an image.

We demonstrate that our MCMS model improves object segmentation performance in electron microscopy images and provides a coherent segmentation of multiple objects. Through speeding up the segmentation process, the proposed method will allow neurobiologists to move beyond individual specimens and analyze populations paving the way for understanding neurodegenerative diseases at the microscopic level.

3.2 M. SEYEDHOSSEINI, M. SAJJADI, AND T. TASDIZEN, "IMAGE SEGMENTATION WITH CASCADED HIERARCHICAL MODELS AND LOGISTIC DISJUNCTIVE NORMAL NETWORKS,"

Contextual information plays an important role in solving vision problems such as image segmentation. However, extracting contextual information and using it in an effective way remains a difficult problem.

To address this challenge, we propose a multi-resolution contextual framework, called cascaded hierarchical model (CHM), which learns contextual information in a hierarchical framework for image segmentation.

At each level of the hierarchy, a classifier is trained based on down sampled input images and outputs of previous levels.

Our model then incorporates the resulting multi-resolution contextual information into a classifier to segment the input image at original resolution.

We repeat this procedure by cascading the hierarchical framework to improve the segmentation accuracy. Multiple classifiers are learned in the CHM; therefore, a fast and accurate classifier is required to make the training tractable. The classifier also needs to be robust against over fitting due to the large number of parameters learned during training.

We introduce a novel classification scheme, called logistic disjunctive normal networks (LDNN), which consists of one adaptive layer of feature detectors implemented by logistic sigmoid functions followed by two fixed layers of logical units that compute conjunctions and disjunctions, respectively.

We demonstrate that LDNN outperforms state-of-the art classifiers and can be used in the CHM to improve object segmentation performance.

3.3 D. LARLUS AND F. JURIE, "COMBINING APPEARANCE MODELS AND MARKOV FIELDS FOR CATEGORY LEVEL OBJECT SEGMENTATION,"

Object models based on bag-of-words representations can achieve state-of-the-art performance for image classification and object localization tasks. However, as they consider objects as loose collections of local patches they fail to accurately locate object boundaries and are not able to produce accurate object segmentation.

On the other hand, Markov random field models used for image segmentation focus on object boundaries but can hardly use the global constraints necessary to deal with object categories whose appearance may vary significantly. In this paper we combine the advantages of both approaches. First, a mechanism based on local regions allows object detection using visual word occurrences and produces a rough image segmentation.

Then, a MRF component gives clean boundaries and enforces label consistency, guided by local image cues (color, texture and edge cues) and by long-distance dependencies. Gibbs sampling is used to infer the model. The proposed method successfully segments object categories with highly varying appearances in the presence of cluttered backgrounds and large view point changes. We show that it outperforms published results on the Pascal VOC 2007 dataset.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

3.4 M. P. KUMAR AND D. KOLLER, "EFFICIENTLY SELECTING REGIONS FOR SCENE UNDERSTANDING,"

Recent advances in scene understanding and related tasks have highlighted the importance of using regions to reason about high-level scene structure.

Typically, the regions are selected beforehand and then an energy function is defined over them. This two step process suffers from the following deficiencies: the regions may not match the boundaries of the scene entities, thereby introducing errors; and as the regions are obtained without any knowledge of the energy function, they may not be suitable for the task at hand.

We address these problems by designing an efficient approach for obtaining the best set of regions in terms of the energy function itself.

Each iteration of our algorithm selects regions from a large dictionary by solving an accurate linear programming relaxation via dual decomposition.

The dictionary of regions is constructed by merging and intersecting segments obtained from multiple bottom-up over segmentations. To demonstrate the usefulness of our algorithm, we consider the task of scene segmentation and show significant improvements over state of the art methods.

3.5 X. HE, R. ZEMEL, AND M. CARREIRA-PERPINAN, "MULTISCALE CONDITIONAL RANDOM FIELDS FOR IMAGE LABELING,"

We propose an approach to include contextual features for labeling images, in which each pixel is assigned to one of a finite set of labels. The features are incorporated into a probabilistic framework, which combines the outputs of several components. Components differ in the information they encode. Some focus on the image-label mapping, while others focus solely on patterns within the label field. Components also differ in their scale, as some focus on fine-resolution patterns while others on coarser, more global structure.

A supervised version of the contrastive divergence algorithm is applied to learn these features from labeled image data. We demonstrate performance on two real-world image databases and compare it to a classifier and a Markov random field.

IV. IMPLEMENTATION

Implementation is the stage of the project when the design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

A. MODULES

- Login modules.
- Fixation-Based Segmentation module.
- Fixated region module.
- Polar Space Method.
- Multiple fixation-based segmentation Module.

4.1 LOGIN MODULES

Login or logon (also called logging in or on and signing in or on) is the process by which individual access to a computer system is controlled by identification of the user using credentials provided by the user.

A user can log in to a system and can then log out or log off (perform a logout / logoff) when the access is no longer needed. Logging out may be done explicitly by the user performing some action, such as entering the appropriate command, or clicking a website link labeled as such. It can also be done implicitly, such as by powering the machine off, closing a web browser window, leaving a website, or not refreshing a webpage within a defined period.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

4.2 FIXATION-BASED SEGMENTATION MODULE

A segmentation framework that takes as its input a fixation (a point location) in the scene and outputs the region containing that fixation.

The fixated region is segmented in terms of the area enclosed by the "optimal" closed boundary around the fixation using the probabilistic boundary edge map of the scene (or image).

The probabilistic boundary edge map, which is generated by using all available visual cues, contains the probability of an edge pixel being at an object (or depth) boundary.

The separation of the cue handling from the actual segmentation step is an important contribution of our work because it makes segmentation of a region independent of types of visual cues that are used to generate the probabilistic boundary edge map.

4.3 FIXATED REGION MODULE

Fixated region is equivalent to finding the "optimal" closed contour around the fixation point. This closed contour should be a connected set of boundary edge pixels (or fragments) in the edge map.

However, the edge map contains both types of edges, namely boundary (or depth) and internal (or texture/intensity) edges.

In order to trace the boundary edge fragments in the edge map to form the closed contour enclosing the fixation point, it is important to be able to differentiate between the boundary edges from the non-boundary (e.g., texture and internal) edges.

4.4 POLAR SPACE METHOD

The optimal contour around the red fixation point on the .disc the gradient edge map of the disc, has two concentric circles. The big circle is the actual boundary of the disc, whereas the small circle is just the internal edge on the disc. The edge map correctly assigns the boundary contour intensity as 0.78 and the internal contour 0.39.

4.5 MULTIPLE FIXATION-BASED SEGMENTATION MODULE

We have described segmentation for a given fixation. Our objective now is to refine that segmentation by making additional fixations inside the initial segmentation to reveal any thin structures not found in the initial segmentation.

Detecting these thin structures can be expensive and complicated if we choose to fixate at every location inside the region.

V. CONCLUSION

We proposed here a novel formulation of segmentation in conjunction with fixation. The framework combines static cues with motion and/or stereo to disambiguate between the internal and the boundary edges. The approach is Motivated by biological vision, and it may have connections to neural models developed for the problem of border ownership in segmentation. Although the framework was developed for an active observer, it applies to image databases as well, where the notion of fixation amounts to selecting an image point which becomes the center of the polar transformation.

Our contribution here was to formulate old problem segmentation in a different way and show that existing computational mechanisms in the state-of-the-art computer vision are sufficient to lead us to promising automatic solutions.

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to all those who provide me the possibility to complete the thesis. I express my sincere thanks to my respected and adored chairman **Mrs.E.Leelavathy Elayappan** and Secretary **Dr.A.K.Ramasamy** for providing support and stimulating environment for developing the thesis work. I am very happy to express my sincere and honorable thanks to **Dr.D.RajaKumari M.Com., M.B.A., M.Phil., Ph.D.,** Principal of our college for her kind encouragement to do this thesis. I would like to

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 12, December 2018

express my sincere thanks to **Mrs.C.Renuga M.Sc(CT)., M.Phil.,** HOD Cum Assistant professor, Department of Computer science and their meticulous guidance, teaching and counseling.

I am deeply indebted to my faculty guide **Mrs.C.Tamilselvi M.Sc.,M.Phil.,B.Ed.,** for this unending support and guidance extended and also my gratitude to all other faculty members of the project committee for their guidance provided to me for the successful completion of this thesis. I also thankful to my family members and friends for their moral support in all part of my work. I extent my thanks to Librarian of Bharathiyar Arts and Science College for Women, for their kind help in providing books and journals for the reference work. Last but not least, Once again, I like to thank Almighty for blessing showered on me.

REFERENCES

1. S. Z. Li, Markov random field modeling in computer vision.Springer Verlag New York, Inc., 1995.
2. C. Farabet, C. Couprie, L. Najman, and Y. LeCun, “Learning hierarchical features for scene labeling,” TdPAMI, 2013.
3. Z. Tu and X. Bai, “Auto-context and its application to high-level vision tasks and 3d brain image segmentation,” TPAMI, vol. 32, no. 10, pp. 1744–1757, 2010.
4. A. Torralba, K. P. Murphy, and W. T. Freeman, “Contextual models for object detection using boosted random fields,” NIPS, 2004.
5. G. Heitz, S. Gould, A. Saxena, and D. Koller, “Cascaded classification models: Combining models for holistic scene understanding, NIPS, 2008.
6. C. Li, A. Kowdle, A. Saxena, and T. Chen, “Toward holistic scen understanding: Feedback enabled cascaded classification models,
7. TPAMI, vol. 34, no. 7, pp. 1394–1408, 2012.
8. E. Jurrus, A. R. C. Paiva, S. Watanabe, J. R. Anderson, B. W. Jones, R. T. Whitaker, E. M. Jorgensen, R. E. Marc, and T. Tasdizen “Detection of neuron membranes in electron microscopy images using a serial neural network architecture,” Medical Image Analysis, vol. 14, no. 6, pp. 770–783, 2010.
9. Z. Ren and G. Shakhnarovich, “Image segmentation by cascaded region agglomeration,” CVPR, 2013.
10. M. Seyedhosseini and T. Tasdizen, “Multi-class multi-scale series contextual model for image segmentation,” Image Processing, IEEE Transactions on, vol. 22, no. 11, pp. 4486–4496, 2013.

BIOGRAPHY

Mrs.K.SARASWATHI, is presently pursuing M.Phil., Final year the Department of Computer Science and Research from Bharathiyar Arts and Science College (w), Deviyakurichi , Attur (tk),Salem(dt) , Tamil Nadu,India[1].

Mrs. Mrs.C.Tamilselvi M.Sc.,M.Phil.,B.Ed., Assistant Professor, Department of Computer Science and Research, Bharathiyar Arts and Science College (w), Deviyakurichi , Attur (tk),Salem(dt), Tamil Nadu,India[2].